# A FRAMEWORK FOR THE DESIGN OF SPEECH-ENABLED SELF-CARE E-HEALTH SYSTEMS

Oyelami, Olufemi Moses[1], Uwadia, Charles Onuwa[2], Akinwale, Adio Taofeek[3]
[1]Covenant University, Ota, Nigeria, 2University of Lagos, Akoka, Nigeria, 3University of Agriculture, Abeokuta, Nigeria
[1]olufemioyelami@gmail.com, olufemioyelam@yahoo.com, [2]couwadia@yahoo.com, [3]atakinwale@yahoo.com

## Abstract

The Internet provides a wide range of health information and services which consumers access for self-care and to participate in a more informed way in their healthcare when they see their physician. This information and services are however, delivered in text form and therefore, does not cater for the needs of the non-computer literate, the visually impaired and the blind. This paper presents a framework for the design of speech-enabled self-care e-Health systems. Reasoning-induced disease diagnosis which existing speech-based disease screening systems lack has been incorporated into the framework to enable systems based on the framework diagnose more than one type of disease. Based on the framework, speech-based self-care e-Health system (SSeS) prototype application was developed. The originality of this framework is that it is speech-based. This takes care of the health needs of the category of people earlier identified and the underserved people, majority of those who are domiciled in Africa.

**Keywords:** Self-Care, Speech-Enabled, VoiceXML, e-Health, Health Information, Reasoning.

## Introduction

Consumers seek health information to educate themselves for self-care and to participate in a more informed way in their healthcare when they see their physicians. As such, they search websites to get this information [1]. This activity is especially helpful in most African countries where healthcare is inadequate [2]. Mobile phone usage is however on the increase on the continent [3].

The information provided by these websites is delivered in text form and this does not cater for the needs of the non-computer literate, the blind as well as the people with visual impairment. Although several frameworks have been proposed for different features in e-Health systems, these frameworks however, do not address the issue of inadequacy of healthcare, visual impairment and non-computer literacy of some sections of the population.

This paper presents a framework for speech-enabled self-care e-Health system. The initiative makes health information and services for self-care available through speech for the categories of people that are underserved and those not taken care of through the graphical user interface (GUI) systems. It incorporates reasoning-induced disease diagnosis which the existing speech-based screening systems lack, to enable them diagnose more than one disease. It also serves as an m-Learning speech-enabled e-Health initiative that provides health information.

## Previous Work

Previous related research efforts in the area of e-Health systems include the following: A framework for the creation, maintenance, and reuse of clinical document templates adhering to HL7 Clinical Document Architecture (CDA) and interoperability with middleware services of the health information infrastructure (HII) is proposed by [4]. The architecture and a prototype of a generic service platform for the provision of mobile healthcare services based on Body Area Networks (BANs) are discussed in [5]. The use of health BANs together with advanced wireless communications enables remote management of chronic conditions and detection of health emergencies whilst maximizing patient mobility. A composite device computing environment (CDCE) framework that can offer access to broad range of multimedia services across multitude of potential devices is proposed in [6]. CDCE needs to be sufficiently adaptive in order to exploit an ever-changing number and diverse range of available computing resources. A mechanism that supports the ubiquitous and efficient exchange of electronic medical records across multiple heterogeneous environments is provided in [7]. Mu-Jung and Mu-Yen proposed a framework for an integrated design of intelligent web-based Chinese Medical Diagnostic System (CMDS) for digestive health using ontology. CMDS used Web interface and expert system to diagnose a number of digestive system diseases [8]. Nsuangani and Pérez presented the result of a study conducted on college students' online activities at health websites. It was found out that those who searched the sites were looking for health-related information, products, and services for self-help [1]. The information provided was however in text form. Medical Library Association listed some websites that people access for self-care, but the access is through GUI only [9]. Sherwani et al.

presented a report of speech-based access to health information by low-literate community health workers to assist them in their job. The information provided was in Urdu language of Pakistan and it was found out that speech technology could help the low-literate [10]. James et al. presented a report of Interactive Voice Response (IVR) developed for screening for early Dementia. A Clinical Dementia Rating Scale was used to assess 156 subjects aged 56 to 93 years. These subjects performed a battery tests administered by an interactive voice response system using standard touch-tone telephones. It was discovered that computer-automated telephone screening using either informant or direct assessment is feasible and that such systems could provide wide-scale, cost-effective screening, education and referral services to patients and care givers [11]. Mundt et al. extended the work presented by James et al. to include educating the callers on dementia disease and making referral services. Data gathered from the pilot testing showed the feasibility of using IVR for education purposes [12]. Kim et al. carried out a test to know the feasibility of using IVR to screen for depression among low-income, urban pregnant patients and to solicit their preferences for treatment. Consenting subjects used a phone in a private clinic room to complete an IVR version of the Edinburgh Postnatal Depression Scale (EPDS). The pilot study suggested the practicability of using an automated phone interview for screening [13]. Corkrey et al. described a pilot interactive voice response cervical cancer screening brief advice interface. The subjects used completed computer-assisted telephone interview (CATI) and the results obtained suggested that an IVR reminder might be useful to increase cervical screening rates [14]. A critical look at the existing speech-based disease screening systems shows that the approach employed by them is that of asking the callers stereotyped set of questions over and over again. There is no reasoning involved in arriving at the conclusion. Because of this limitation, they are mainly meant for diagnosing a particular type of disease and consequently cannot handle more than one disease.

## Materials and Methods

## Methodology

In developing the framework, the American Medical Library Association recommended websites were accessed to gather the features provided by the framework. Reasoning-induced disease diagnosis which existing speech-based disease screening systems lack has been provided as part of the features to enable systems based on the framework diagnose more than one disease. A prototype voice application was developed to provide functionalities based on the features of the framework using four types of fever prevalent in Africa as case diseases. VoiceObjects Desktop for Eclipse was used to develop, test and deploy the prototype application. Voxeo Prophecy was used as the media platform, and server-side Javascript was used to specify the rules with which the system reasons in diagnosing a disease using script objects. Logic objects were used to control the logic of the dialog flow. VoiceObjects embedded database was used to store the context of the problem domain, while the inbuilt rule engine within VoiceObjects determines which rule gets fired. X-Lite soft phone was used to call the application. These tools were chosen because they allow a developer to develop and test an application locally, eliminating the need to pick up a phone and dial a specific number each time the application is to be tested and debugged.

## Speech Technology

Speech technology is a technology that is capable of catering for the needs of the categories of people that are blind, visually impaired and non-computer literate. Until recently, Internet applications have primarily been dependent on visual interfaces to provide access to information and services. Advances in speech recognition technology are however, allowing the creation of voice applications that users interact with by speaking to them through telephones rather than by using traditional input devices. All that is necessary for them is to obtain the information they require through voice. Driving this technology is Voice Extensible Markup Language or VoiceXML. VoiceXML is a standard language for building interfaces between voice recognition software and Web content. Just as HTML defines the display and delivery of text and images on the Internet, VoiceXML translates XML-tagged Web content into a format that speech recognition software can deliver by phone [15].

VoiceXML technology follows the same model as the HTML and Web browser technologies. Similar to HTML, a VoiceXML application does not contain any platform specific knowledge for processing the content. The platform specific processing capability is provided through the Voice Gateway that incorporates Automatic Speech Recognition (ASR) and Text-to-Speech (TTS) engines, along with a Voice Browser [17]. The Voice Gateway architecture is presented in figure 1. It uses the familiar client – server paradigm.

VoiceXML pages are served from a Web Server and is typically created dynamically within the framework of an Application Server. The Voice Browser is a telephone-based browser that renders VoiceXML pages audibly. The users interact with the Automatic Speech Recognition (ASR) engine and Text To Speech (TTS) voice synthesis and restitution of pre-recorded audio files are the outputs the machine can use.

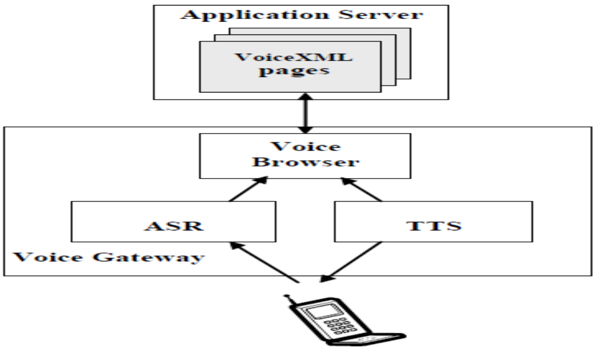


Figure 1. Voice gateway architecture (Source : 16)

## e-Health

e-Health is a term that is commonly used in relation to the use of technology in healthcare. Although there is no universally acceptable meaning of e-Health [17], it could be said to be the use of Information and Communication Technologies (ICT) to improve health and healthcare systems [18]. e-Health combines the use of electronic communication and information technology in the health sector, or is the use of digitally transmitted data that is stored and retrieved electronically for clinical, educational and administrative purposes both at a local site and remotely in the health sector [19].

## m-Learning and e-Learning

According to Niall Winters [21], current perspectives on mobile learning generally fall into the following four broad categories:

(1) Technocentric: Here, mobile learning is viewed as learning using a mobile device such as PDA, mobile phone, iPod, playstation portable, etc.

(2) Relationship to e-learning: This perspective characterizes mobile learning as an extension of e-learning.

(3) Augmenting formal education: A means of augmenting formal education.

(4) Learner-centered: Any sort of learning that happens when the learner is not at a fixed, predetermined location or learning that happens when the learner takes advantage of learning opportunities offered by mobile technologies. According to Parsons and Ryu , m-learning is broadly defined as the delivery of learning content to learners utilizing mobile computing devices [21], and according to Kambourakis G. et al., m-learning is the point at which mobile computing and e-learning intersect to produce an anytime, anywhere learning experience [22]. From the various definitions of m-learning, we can conclude that m-learning is a form of e-learning that involves any learning with the use of mobile device to produce an anywhere and anytime learning experience to cater for the needs of different learners and augmenting their formal learning experience.

On the other hand, there are various definitions of e-learning and one thing that is common to them all is the term ‘technology’. It involves a learner using a variety of computer and networking technologies to access training materials. It is education offered using electronic delivery methods such as CD-ROMs, video conferencing, websites and e-mails to facilitate and enhance learning through the use of devices based on computer and communications technology [23]. It covers a wide set of applications and processes such as Web-based learning, computer-based learning (CD-ROM), virtual classrooms, and digital collaboration [24]. In an overall definition, it is the use of electronic technology to support and enhance teaching and learning.

## The Architectural Framework for the Speech-Enabled Self-Care e-Health System (SSeS)

Figure 2 below shows the architecture for the proposed framework.

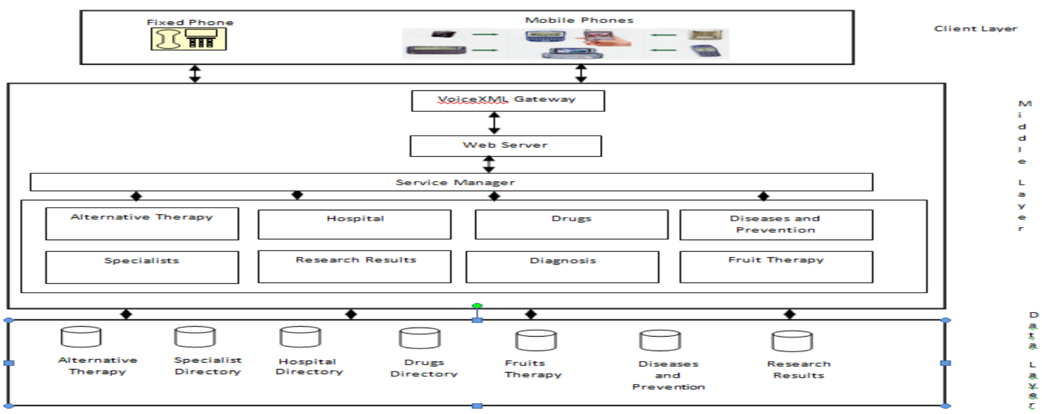


Figure 2. The architecture for the framework

The architecture consists of the client layer, the middle layer and the data layer. The client layer consists of telephone either fixed or mobile with which the caller will interact with the middle layer. In the middle layer, the VoiceXML gateway is an operational platform on which VoiceXML services run. As a gateway, it bridges the world of telephony to other networks such as Internet and data networks. The Web server serves the VoiceXML pages and the service manager allows the caller to pick any of the services below:

**Alternative Therapy:** This provides alternative therapy to any disease the caller wants information about. This is especially useful in the context of some African nations like Nigeria where the National Agency for Food and Drug Administration and Control (NAFDAC) has registered some alternative therapy medicines.

**Hospital:** This service provides a list of hospitals in the caller's proximity in case of the need to consult a doctor and in cases of emergency.

**Drugs:** This provides over-the-counter (OTC) drugs for specific diseases, which the caller can purchase for treatment.

**Diseases and Prevention:** The service provides information about diseases and how they can be prevented.

**Research Results:** It provides information about new results of researches related to treatment, prevention, how to live a healthy life, etc.

**Specialists:** This provides information about specialist doctors in the proximity of the caller in case of the need to consult one.

**Diagnosis:** This service allows the caller to diagnose the kind of disease he/she is suffering from.

**Fruit Therapy:** This provides information about fruits that can be used to treat and prevent certain diseases. This is especially useful for people in the rural areas in Africa where healthcare is inadequate and where the occupation of the entire populace is farming.

The data layer consists of the database that houses the information accessible through the middle layer including all the words and phrases which the user is expected to say.

## Speech-Based Self-Care e-Health System (SSeS) System Design

Based on the framework provided above, Speech-Based Self-Care e-Health System (SSeS) prototype was developed. Figure 3 shows the control of flow diagram for SSeS.

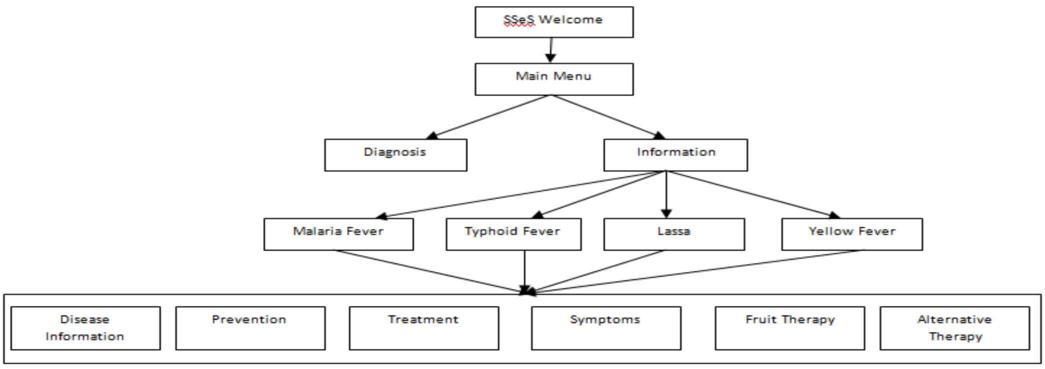


Figure 3. SSeS control flow diagram

The system has a welcome message played to intimate the caller of the services it renders and the insruction on how to use the system. The caller is then given a choice of either diagnosing his/her fever or getting information about a select list of fever rampant in Africa (malaria fever, typhoid fever, lass fever and yellow fever). For each of the fever, the caller can get general information about the fever (e.g how it can be contacted, where it is prevalent, etc), how the fever can be prevented, how it can be treated using orthodox medicine, the symptoms of the fever, how the fever can be treated with fruits and alternative therapies available for the treatment of the fever. The system is able to diagnose more than one disease because of the reasoning introduced into it. This approach is an improvement over the existing speech-based disease screening systems that can only handle only one disease.

## Reasoning Model of Speech-Based Self-Care e-Health System (SSeS)

The table below shows the symptoms associated with each of the four types of fever diagnosed by SSeS.

Table 1: Symptoms of the diagnosed fevers

| **Malaria Fever** | **Lassa Fever** | **Typhoid Fever** | **Yellow Fever** |
|---|---|---|---|
| Headache | Headache | Headache | |
| Stomach pain | Stomach pain | Stomach pain | Stomach pain |
| Fever | Fever | Fever | Fever |
| Diarrhea | Diarrhea | Diarrhea | |
| Vomiting | Vomiting | | Vomiting |
| Chill | Chill | | Chill |
| Nausea | | | Nausea |
| | | | |
| | Back pain | | Back pain |
| | Fatigue | Fatigue | Fatigue |
| | Muscle pain | | Muscle pain |
| | Loss of appetite | Loss of appetite | |
| | | | |
| Bitter taste | Chest pain | Rash | Bone pain |
| Sweating | Cough | Bloody stool | Bleeding |
| Joint pain | Swollen eyes | | Red eyes |
| | Swollen neck | | Low urine |
| | Swollen face | | Blood vomiting |
| | Ringing in ears | | |

Table 1 is divided into three sections. The symptoms from the first two sections are not sufficient to diagnose any of the four types of fever because the different types of fever have some of them in common. However, the symptoms from the third section, which is the last, are sufficient to diagnose each of the fever types because they are specific to the different types of fever. For instance, any of bitter taste, sweating and joint pain is enough to suggest malaria fever. If the caller mentions any three symptoms from any of the first, second or the first two sections, the system must carry out further interrogations to ascertain the kind of fever the caller is suffering from. The algorithm below shows the reasoning model used by SSeS in diagnosing the particular type of fever a caller is suffering from:

**Begin**

1. Capture three symptoms from the caller.
2. If any of the symptoms is from bitter taste, sweating and joint pain, report malaria fever as the fever diagnosed.
3. Else if any of the symptoms is either rash or bloody stool, report typhoid fever as the fever diagnosed.
4. Else if any of the symptoms is from bone pain, bleeding, red eyes, low urine and blood vomiting, report yellow fever as the fever diagnosed.
5. Else if any of the symptoms is from chest pain, cough, swollen eyes, swollen neck swollen face and ringing in ears, report lassa fever as the fever diagnosed.
6. // Symptoms supplied by the caller are common to two or more of the fever types
   Else interrogate the caller until the fever type is diagnosed.

**End**

In the algorithm above, the caller is expected to mention three symptoms through which SSeS will determine the nature of the fever. If the caller mentions any symptom(s) specific to a particular kind of fever, the disease is diagnosed and reported to the caller. If however, the symptoms mentioned are common to two or more of the fever types, SSeS will ask the caller questions in order to be able to determine the particular kind of fever.

Figure 4 below shows SSeS being called using X-Lite soft phone.


Figure 4. Calling SSeS using X-Lite soft phone

## Conclusion

This paper has proposed a framework for the design of speech-enabled self-care e-Health systems to make health information and services available through speech and accessed by telephone. This will empower for self-care, the estimated 180, 000,000 visually impaired and blind people worldwide [25], the non-computer literate, as well as the underserved people, majority of those who reside in places in Africa where healthcare is inadequate. The initiative will also aid their efforts in seeking healthcare. Reasoning-induced disease diagnosis which has been incorporated into the framework enables systems based on the framework to diagnose more than one disease. This is an improvement over the existing speech-based disease screening systems which can only diagnose one disease.

## References


[1] Nsuangan, N. M. and Pérez, M. A. Accessing Web-Based Health Related Information by College Students: An Exploratory Study, Californian Journal of Health Promotion, 4(1), 2006, pp. 64-74.

[2] World Health Organization. A Safer Future: Global Public Health Security in the 21st Century, World Health Orgaization: Geneva, Switzerland, 2007.

[3] Euromonitor International, No end in sight for Africa's mobile telephone boom, 2007. Available from http://www.euromonitor. com No_end_in_sight_for_Africas_mobile_ telephone_boom. [April, 2008]

[4] Chronaki, C. E., Lelis, P., Demou, C., Tsiknakis, M. and Orphanoudakis, S. C. An HL7/CDA Framework for the Design and Deployment of Telemedicine Services, Proceedings of IEEE EMBC, Instambul, Turkey, 2001, pp. 25-28.

[5] Wac, J., Bults, R., Konstantas, D., Van, A. and Halteren, V. Mobile Health Care Over 3G Networks:The Mobihealth Pilot System and Service, Proceedings of the Gobal Mobile Congress, GMC, 2004. Available from http://cui.unige.ch/wac/ publications/ WacGlobalMobileCongress_2004 pdf [June, 2009]

[6] Pham, T., Schneider, G. and Goose, S. Composite Device Computing Environment: A Framework for Augmenting PDA Using Surrounding Resources, Proceedings of Handheld and Ubiquitous Computing, Springer, 2000.

[7] Kallepalli, V. N. A Security Framework for DICOM Images in Health, Information Systems. M. Sc. Thesis, Department of Computer Science, Faculty of Graduate Studies, University of Manitoba, Canada, 2003.

[8] Mu-Jung, H. and Mu-Yen, C. Integrated design of the intelligent web-based Chinese Medical Diagnostic System (CMDS) – Systematic development for digestive health, Expert System with Applications, 32, 2007, pp. 658-673.

[9] Medical Library Association (MLA), A User's Guide to Finding and Evaluating Health Information on the Web, 2008. Available from http://www.mlanet.org/resources/userguide.html [May, 2008]

[10] Sherwani, J., Ali, N., Mirza, S., Fatma, A., Memon, Y., Karim, M., Tongia, R., Rosenfeld, R. HealthLine: Speech-based Access to Health Information by Low-literate Users. Proceedings of IEEE/ACM Int'l Conference on Information and Communication Technologies and Development, Bangalore, India, 2007. Available from http://www.cs.cmu.edu/~jsherwan/pubs/ictd07.pdf [August 2008]

[11] James, C., Mundt, K. L., Ferber, M. R. and John, H. Computer–Automated Dementia Screening Using a Touch-Tone Telephone. ARCH INTERN MED, 161, 2001, pp. 2481-2487. Available from http://www.cs.cmu.edu/~jsherwan/pubs/ictd07.pdf [25 April 2008]

[12] Mundt, J. C., Moore, H. K. and Greist, J. H. A Novel Interactive Voice Response (IVR) System for Dementia Screening, Education and Referral: One-Year Summary. Alzheimer Disease & Associated Disorder, 19 (3), 2005, pp.143-147.

[13] Kim, H., Bracha, Y. and Tipnis, Automated Depression Screening in Disadvantaged Pregnant Women in an Urban Obstetric Clinic, Archives of Women's Mental Health, 10 (4), August 2007, pp. 163-169.

[14] Corkrey, R., Parkinson, L., Bates, L., Green, S. and Htun, A. T. Pilot of a Novel Cervical Screening Intervention: Interactive Voice Response, Australian and New Zealand Journal of Public Health, 29 (3), 2005, pp. 261-264.

[15] Stephen, K., Kevin, C., Rob, M. and Mark, S. A VoiceXML Enabled Locus of control System Designed to Engage School Pupils Lacking Commitment in the career development and decision making process. Georgian Electronic Scientific Journal: Education Science and Psychology, 1(10), 2007, pp. 45-53.

[16] Kondratova, I. Voice and Multimodal Technology for the Mobile Worker, ITcon 9, 2004, pp. 345-353.

[17] Hans, O., Rizo, C., Enkin, M. and Jadad, A. "What is eHealth (3): A Systematic Review of Published definitions", Journal of Medical Internet Research, 7(1), 2005.

[18] Alvarez R. C. The Promise of e-health- a canadian perspective. eHealth Int 1 ( 4), 2002.

[19] Ontario Hospital eHealth Council, An eHealth Blueprint. Setting the Course for Action. An Ontario Hospital perspective, 2001. Available from http://www.oha.com/ [June, 2009]

[20] Sharples, M., Big issues in mobile learning: Report of a workshop by the kaleidoscope network of excellence, mobile learning initiative, 2006. Available from http://telearn.noe-kaleidoscope.org/open-archive/file? Sharples_Big_Issues.pdf. [November, 2007]

[21] Parsons, D. and Ryu, H., A framework for assessing the quality of mobile learning, 2006. Available from http://www. irrodl.org/index.php/irrodl/ article/view/348/873 [November, 2007]

[22] Kambourakis, G., Kontoni, D.P.N. and Sapounas, I., Introducing attribute certificates to secure distributed e-learning or m-learning services. IAS

TED International Conference. Innsbruck, Australia, 2004, pp. 436-440, Available from http://www.ice.upc.edu/butlleti/inns bruck/416-174.pdf. [ May, 2007]
[23] Hobsons MBA Service, 2005 http://www.mba.hobsons.com/glossary.jsp
[24] Glossary of Terms, 2002. Available from http://www.neiu.edu/~dbehrlic/hrd408/glossary.htm. [May, 2008]
[25] World Health Organization, Global initiative for the prevention of avoidable blindness. World Health Organization, Geneva, Switzerland. 1997.